

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 08-152900
(43)Date of publication of application : 11.06.1996

(51)Int.Cl. G10L 5/04
G10L 9/18

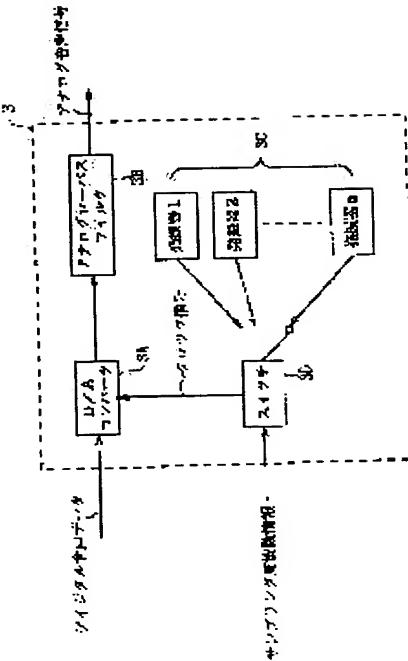
(21)Application number : 06-319217 (71)Applicant : SONY CORP
(22)Date of filing : 28.11.1994 (72)Inventor : IWAHASHI NAOTO
YAMADA KEIICHI

(54) METHOD AND DEVICE FOR VOICE SYNTHESIS

(57)Abstract:

PURPOSE: To synthesize a voice having a voice quality which differs from the voice quality of the voice data that are beforehand stored.

CONSTITUTION: The sampling frequency of a digital-to-analog converter 3, which converts digital voice data into analog voice data, is converted by sampling frequency converting means 3A, 3B, 3C and 3D and the digital voice data are converted into analog voice data by the converted sampling frequency. Thus, a synthesized voice having a good voice quality, a synthesized voice having a desired voice quality and synthesized voices having various voice qualities are easily and surely outputted. Furthermore, voices having different voice qualities are stably synthesized.



LEGAL STATUS

[Date of request for examination] 28.11.2001

[Date of sending the examiner's decision of rejection] 05.12.2003

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

THIS PAGE BLANK (USPTO)

1

【特許請求の範囲】

【請求項1】デジタル記録された音声データより音声を合成する音声合成方法において、合成文の内容を表す合成文情報に基づいて上記デジタル音声データを合成し、当該デジタル音声データをアナログ音声信号に変換するデジタルアナログ変換器のサンプリング周波数を変換し、当該変換されたサンプリング周波数により上記デジタル音声データを上記アナログ音声信号に変換することを特徴とする音声合成方法。

【請求項2】予め蓄積しておいた音声データとして男性が発声した音声データを用いて、上記デジタルアナログ変換器のサンプリング周波数を、上記音声データを蓄積する際に用いたサンプリング周波数よりも高く設定することを特徴とする請求項1に記載の音声合成方法。

【請求項3】上記デジタルアナログ変換器のサンプリング周波数を変換する場合、当該サンプリング周波数の変化に応じて合成音声の基本周波数及び音韻継続時間を設定することを特徴とする請求項1に記載の音声合成方法。

【請求項4】デジタル記録された音声データより音声を合成する音声合成方法において、合成文の内容を表す合成文情報に基づいて上記デジタル音声データを合成し、当該デジタル音声データを再サンプリングして上記デジタル音声データのサンプリング周波数を変換することを特徴とする音声合成方法。

【請求項5】予め蓄積しておいた音声データとして男性が発声した音声データを用いて、上記デジタル音声データを再サンプリングする際のサンプリング周波数を、上記音声データを蓄積する際に用いたサンプリング周波数よりも低く設定することを特徴とする請求項4に記載の音声合成方法。

【請求項6】上記デジタル音声データを再サンプリングして当該デジタル音声データのサンプリング周波数を変換する場合、当該サンプリング周波数の変化に応じて合成音声の基本周波数及び音韻継続時間を設定することを特徴とする請求項4に記載の音声合成方法。

【請求項7】デジタル記録された音声データより音声を合成する音声合成装置において、合成文の内容を表す合成文情報に基づいて上記デジタル音声データを合成するデジタル音声データ合成手段と、

上記デジタル音声データをアナログ音声信号に変換するデジタルアナログ変換器のサンプリング周波数を変換するサンプリング周波数変換手段とを具え、当該サンプリング周波数変換手段で変換したサンプリング周波数により上記デジタル音声データを上記アナログ音声信号に変換することを特徴とする音声合成装置。

10

【請求項8】デジタル記録された音声データより音声を合成する音声合成装置において、合成文の内容を表す合成文情報に基づいて上記デジタル音声データを合成するデジタル音声データ合成手段と、上記デジタル音声データを再サンプリングして上記デジタル音声データのサンプリング周波数を変換するサンプリング周波数変換手段とを具えることを特徴とする音声合成装置。

【発明の詳細な説明】

【0001】

【目次】以下の順序で本発明を説明する。

産業上の利用分野

従来の技術

発明が解決しようとする課題

課題を解決するための手段（図1、図2及び図4）

作用（図2及び図4）

実施例

（1）第1実施例による音声合成装置（図1～図3）

20 （2）第2実施例による音声合成装置（図4及び図5）

発明の効果

【0002】

【産業上の利用分野】本発明は音声合成方法及び音声合成装置に関し、例えばカーナビゲーションシステムの音声案内機能、電子化された文書の読み上げシステムや電子ゲーム機器などに適用して好適なものである。

【0003】

【従来の技術】従来、音声の声質を特徴づける要因の1つとして、音声のスペクトル包絡の特性があげられる。

30 スペクトル包絡は、発生器官の共振、反共振特性、声帯音源のスペクトル概形、口唇や鼻孔などからの音波の放射特性などを含めたスペクトルの大局的な特徴を表す。このスペクトル包絡の特性は、主に発声器官の物理的な形状に依存しており、スペクトル特性を変えることによつて声質を変化させることができる。

【0004】

【発明が解決しようとする課題】ところで声質を変化させるためのスペクトル特性変換法として、スペクトルの極周波数を変換する方式（箱田 和雄「極制御による男女声質変換法」日本音響学会講演論文集、213頁～214頁、1987年10月）が提案されている。この方式では、まず音声スペクトルを全極型のスペクトル特性で近似、すなわち複数の極周波数とそれに関するバンド幅で表す。この極周波数を決められた規則に従つて変換することによりスペクトルの特性を変える。こうして得られたスペクトル特性を有するデジタルフィルタを構成し、このフィルタを用いて声質の異なる音声を合成している。

【0005】ところがこの方式では、極周波数の変換によつて、得られたスペクトル特性を有するデジタルフィルタを構成しなければならず、そのための演算が複雑

40

40

50

なるという問題があつた。

【0006】本発明は以上の点を考慮してなされたもので、所望の声質の合成音声を出し得る音声合成方法及び音声合成装置を提案しようとするものである。

【0007】

【課題を解決するための手段】かかる課題を解決するため本発明においては、デジタル記録された音声データより音声を合成する音声合成方法において、合成文の内容を表す合成文情報に基づいてデジタル音声データを合成し、当該デジタル音声データをアナログ音声信号に変換するデジタルアナログ変換器（3）のサンプリング周波数を変換し、当該変換されたサンプリング周波数によりデジタル音声データをアナログ音声信号に変換する。

【0008】また本発明においては、デジタル記録された音声データより音声を合成する音声合成方法において、合成文の内容を表す合成文情報に基づいてデジタル音声データを合成し、当該デジタル音声データを再サンプリングしてデジタル音声データのサンプリング周波数を変換する。

【0009】また本発明においては、デジタル記録された音声データより音声を合成する音声合成装置（1）において、合成文の内容を表す合成文情報に基づいてデジタル音声データを合成するデジタル音声データ合成手段（2）と、デジタル音声データをアナログ音声信号に変換するデジタルアナログ変換器（3）のサンプリング周波数を変換するサンプリング周波数変換手段（（3A）、（3B）、（3C）、（3D））とを設け、当該サンプリング周波数変換手段（（3A）、（3B）、（3C）、（3D））で変換したサンプリング周波数によりデジタル音声データをアナログ音声信号に変換する。

【0010】また本発明においては、デジタル記録された音声データより音声を合成する音声合成装置（1）において、合成文の内容を表す合成文情報に基づいて上記デジタル音声データを合成するデジタル音声データ合成手段（2）と、デジタル音声データを再サンプリングしてデジタル音声データのサンプリング周波数を変換するサンプリング周波数変換手段（12）とを設ける。

【0011】

【作用】デジタル音声データをアナログ音声信号に変換するデジタルアナログ変換器（3）のサンプリング周波数をサンプリング周波数変換手段（（3A）、（3B）、（3C）、（3D））によって変換し、当該変換されたサンプリング周波数によりデジタル音声データをアナログ音声信号に変換する。これにより、良好な合成音声、所望の声質の合成音声、様々な声質の合成音声を簡易かつ確実に出力することができると共に、声質の異なる音声を安定して合成することができる。

【0012】またデジタル音声データをサンプリング周波数変換手段（12）によって再サンプリングして当該デジタル音声データのサンプリング周波数を変換する。これにより、良好な合成音声、所望の声質の合成音声、様々な声質の合成音声を簡易かつ確実に出力することができると共に、声質の異なる音声を安定して合成することができる。

【0013】

【実施例】以下図面について、本発明の一実施例を詳述する。

【0014】（1）第1実施例による音声合成装置
図1において、1は全体として本発明の第1実施例による音声合成装置の概略構成を示している。音声合成装置1はデジタル音声データ合成器2、デジタルアナログ変換器（D/A変換器）3及びスピーカ4より構成されており、D/A変換器3のサンプリング周波数は可変となるように構成されている。この音声合成装置1では、D/A変換器3のサンプリング周波数をサンプリング周波数情報に応じたサンプリング周波数に変換して出力音声の声質を変化させている。

【0015】図1に示すように、デジタル音声データ合成器2には、合成しようとする文の情報（合成文情報）が入力される。合成文情報としては、例えば合成文の内容を表す [kyo' owa te' nkiga i' i desune] のような音韻系列情報及びアクセント情報である。ここで「」はアクセントの位置を表す。デジタル音声データ合成器2では、合成文情報に基づいてデジタル音声データを合成し、このデジタル音声データをD/A変換器3に送出する。

【0016】D/A変換器3では、サンプリング周波数情報によって指定されたサンプリング周波数によりデジタル音声データをアナログ音声信号に変換する。D/A変換器3より出力されたアナログ音声信号は音声としてスピーカ4より出力される。

【0017】D/A変換器3の構成を図2に示す。図2に示すように、D/A変換器3は、デジタル音声データをアナログ音声信号に変換するデジタルアナログコンバータ（D/Aコンバータ）3Aと、D/Aコンバータ3Aより出力されるアナログ音声信号に含まれる高周波数帯域の信号成分を除去するアナログローパスフィルタ3Bと、サンプリング周波数情報に基づいて発振器3Cを選択するスイッチ3Dと、それぞれが異なる分周率を有する複数（1～n）の発振器3Cとによって構成されている。

【0018】以上の構成において、D/A変換器3に入力されたデジタル音声データは、まずD/Aコンバータ3Aにおいてアナログ音声信号に変換される。このときD/Aコンバータ3Aのサンプリング周波数はスイッチ3Dによって選択された発振器3Cで生成されるクロノツク信号により変更され、変更されたサンプリング周波

数によりデジタル音声データをアナログ音声信号に変換する。

【0019】すなわちスイッチ3Dにはサンプリング周波数情報が入力され、スイッチ3Dはこのサンプリング周波数情報に基づいて発振器3Cを選択し、選択された発振器3Cで生成されるクロツク信号がD/Aコンバータ3Aに出力される。従つてD/Aコンバータ3Aのサンプリング周波数はこのクロツク信号に基づいて変更される。D/Aコンバータ3Aより出力されたアナログ音声信号はアナログローパスフィルタ3Bにおいて、当該信号に含まれている高周波数帯域の信号成分が除去されてアナログ音声信号としてスピーカ4に出力される。

【0020】ここで従来の音声合成装置では、D/A変換器のサンプリング周波数と、蓄積してある音声をA/D（アナログ/デジタル）変換した時のサンプリング周波数とはほぼ等しい。ただしシステム構成上若干の違いがある場合もある。

【0021】図3にサンプリング周波数を変更する前の音声スペクトルの変化（A）とサンプリング周波数を1.2倍に変化させたときの音声スペクトルの変化（B）を示す。D/A変換器3でサンプリング周波数を変えた場合、変換倍率が「1」より大きいときはスペクトル全体が高域側に移動し、変換倍率が「1」より小さいときは低域側に移動する。

【0022】以上の構成によれば、合成音声の出力に用いるD/A変換器3のサンプリング周波数を、サンプリング周波数情報に基づいて選択した発振器3Cで生成されるクロツク信号により変更し、この変更したサンプリング周波数によってデジタル音声データをアナログ音声信号に変換したことにより、音声合成装置1に予め蓄積しておいた音声データの声質とは異なる声質を有する音声を容易に合成することができる。

【0023】例えば、男性の声データを予め音声合成装置1に蓄積しておいた場合、D/A変換器3のサンプリング周波数を、蓄積しておいた男性の声データの元のサンプリング周波数よりも高くなるように変更することにより、女性らしい声質の合成音声を出力することができる。

【0024】また上述の構成によれば、D/A変換器3のサンプリング周波数を変更しても、音声のスペクトル構造を特徴づける「ホルマント（formant）」と呼ばれる局所的なピークの相対的な構造に変化を生じさせないので、非音声的にならず安定して声質を変換することができる。一般に女性の音声スペクトルのホルマント周波数は男性の音声スペクトルのホルマント周波数よりも高いので、男性の音声データを変換して女性らしく聞こえる音声スペクトルを得るために、サンプリング周波数を高くすれば良い。

【0025】（2）第2実施例による音声合成装置

図1との対応部分に同一符号を付して示す図4において

て、10は全体として本発明の第2実施例による音声合成装置を示している。この音声合成装置10では、デジタル音声データ合成器2とD/A変換器11との間にサンプリング周波数変換演算器12が設けられている。このサンプリング周波数変換演算器12によってデジタル音声データのサンプリング周波数をサンプリング周波数情報に応じて再サンプリングすることによりデジタル音声データのサンプリング周波数を変更する。このサンプリング周波数変換演算器12は、例えばデジタルシグナルプロセッサ（digital signal processor, DSP）より構成されている。

【0026】サンプリング周波数変換演算器12における演算処理の流れを図5に示す。ここでデジタル音声データ合成器2より出力されるデジタル音声データのサンプリング周波数をX [Hz]とする。デジタル音声データ合成器2より出力されたデジタル音声データをアップサンプリングしてサンプリング周波数をm倍のmX [Hz]とする。次にこのm倍されたサンプリング周波数を1/n倍にダウンサンプリングして(m/n)X [Hz]のサンプリング周波数を有するデジタル音声データを得る。ここでm、nは自然数である。

【0027】(m/n)X [Hz]のサンプリング周波数に変換されたデジタル音声データはD/A変換器11でアナログ音声信号に変換された後、音声としてスピーカ4より出力される。このようにサンプリング周波数変換演算器12によってデジタル音声データのサンプリング周波数を変換した場合、変換倍率が「1」より大きいときはスペクトル全体が高域側に移動し、変換倍率が「1」より小さいときは低域側に移動する。すなわちD/A変換器3でサンプリング周波数を変換した場合と逆の結果になる。

【0028】以上の構成によれば、デジタル音声データ合成器2とD/A変換器11との間にDSPで構成されるサンプリング周波数変換演算器12を設け、このサンプリング周波数変換演算器12でデジタル音声データのサンプリング周波数をn/m倍に変換したことにより、音声合成装置10に予め蓄積しておいた音声データの声質とは異なる声質を有する音声を容易に合成することができる。

【0029】また上述の構成によれば、サンプリング周波数変換演算器12でデジタル音声データのサンプリング周波数をm/n倍に変換しても、音声のスペクトル構造を特徴づける「ホルマント」と呼ばれる局所的なピークの相対的な構造に変化を生じさせないので、非音声的にならず安定して声質を変換することができる。また上述の構成によれば、サンプリング周波数変換演算器12をDSPで構成したことにより、合成音声の品質を保持したまま少ない演算量で声質を変換することができる。

【0030】ここで第1実施例及び第2実施例で述べた

ようにサンプリング周波数をサンプリング周波数情報に基づいて変換することによつてスペクトルの特性を変えた場合、同時に音韻継続時間や基本周波数も変化する。従つて上述の第1及び第2実施例では、出力される音声が適切になるように基本周波数や音韻継続時間などのいわゆる韻律情報をサンプリング周波数の変更に応じて適切に設定する。この場合、スペクトル特性と、音韻継続時間や基本周波数とをそれぞれ独立に変化させる。

【0031】すわなちD/A変換器3のサンプリング周波数を変えるか又はサンプリング周波数変換演算器12でデジタル音声データのサンプリング周波数を変えることによつて出力音声の声質を変化させる場合、出力される音声が適切なものとなるように、予めサンプリング周波数の変化を補正するように音韻継続時間や基本周波数を決定する。

【0032】例えばD/A変換器3のサンプリング周波数を1.2倍に変換した場合、得られる合成音声の音韻継続時間は元の1/1.2倍となり、早口な合成音声になる。従つて元の音声の音韻継続時間をD/A変換器3のサンプリング周波数に合わせて1.2倍にして適切な発声速度を得る。また基本周波数に関しても、変換後は1.2倍になる。出力したい音声の基本周波数がY [Hz]であるとき、変換前の合成音の基本周波数をY/1.2 [Hz]となるように音声を合成する。このように音韻継続時間や基本周波数を変える合成方式に関しては、例えば「デジタル音声処理」古井著、東海大学出版会(1985)などに詳述されている。

【0033】なお上述の実施例においては、スイッチ3Dによつて発振器3Cを選択し、選択した発振器3Cで生成されるクロツク信号によつてD/Aコンバータ3Aのサンプリング周波数を変換した場合について述べたが、本発明はこれに限らず、D/Aコンバータ3Aにクロツク信号を送出しないスイッチモードを設け、声質を変換せずに合成音声を出力してもよい。

【0034】また上述の実施例においては、デジタル音声データのサンプリング周波数を変換するサンプリング周波数変換演算器12をDSPで構成した場合について述べたが、本発明はこれに限らず、他の演算装置で構成してもよい。

【0035】

【発明の効果】上述のように本発明によれば、デジタル音声データをアナログ音声信号に変換するデジタルアナログ変換器のサンプリング周波数をサンプリング周波数変換手段によつて変換し、当該変換したサンプリング周波数によりデジタル音声データをアナログ音声信号に変換することにより、良好な合成音声、所望の声質の合成音声、様々な声質の合成音声を簡易かつ確実に出力することができる共に、声質の異なる音声を安定して合成することができる。

【図面の簡単な説明】

【図1】第1実施例による音声合成装置の構成を示すプロツク図である。

【図2】D/A変換器の構成を示すプロツク図である。

【図3】サンプリング周波数を変換する前の音声スペクトルの変化(A)とD/A変換器によつてサンプリング周波数を変換した後の音声スペクトルの変化(B)との説明に供する周波数特性曲線図である。

【図4】第2実施例による音声合成装置の構成を示すプロツク図である。

【図5】サンプリング周波数変換演算器によるサンプリング周波数の変換処理の流れを示すプロツク図である。

【符号の説明】

1、10……音声合成装置、2……デジタル音声データ合成器、3、11……D/A変換器、3A……D/Aコンバータ、3B……アナログローパスフィルタ、3C……発振器、3D……スイッチ、4……スピーカ、12……サンプリング周波数変換演算器。

【図5】

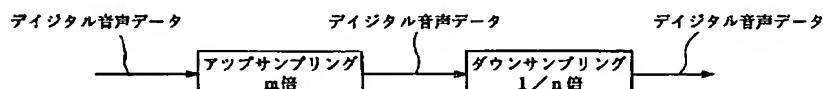


図5 デジタル音声データのサンプリング周波数の変換処理

【図1】

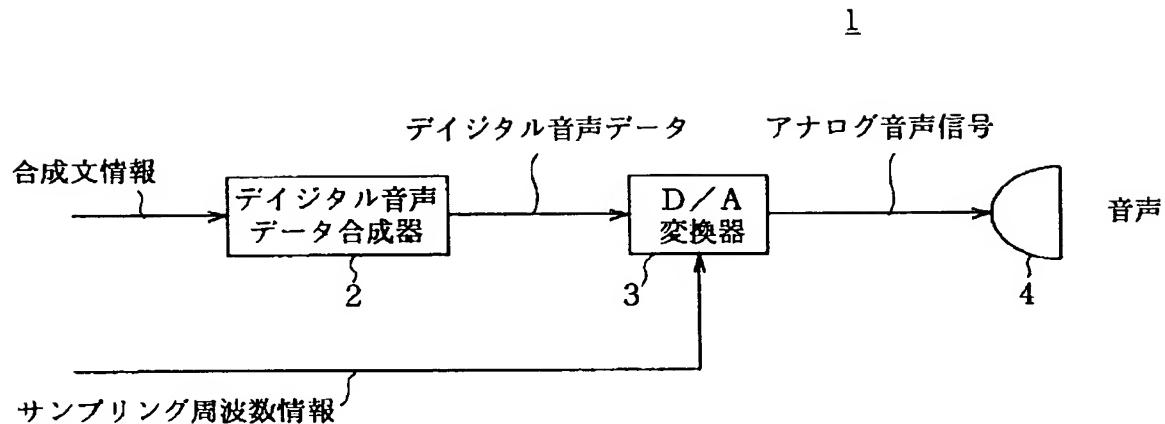


図1 第1実施例による音声合成装置の構成

【図2】

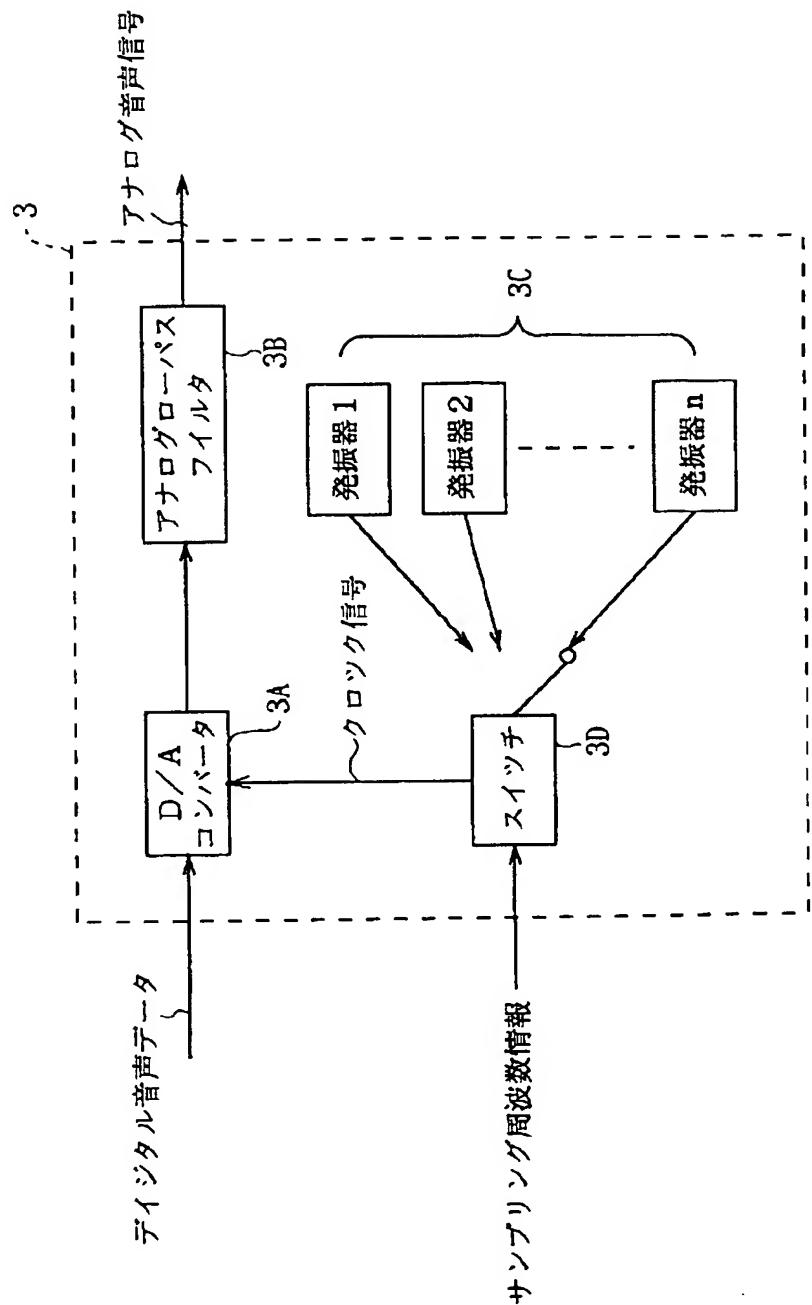
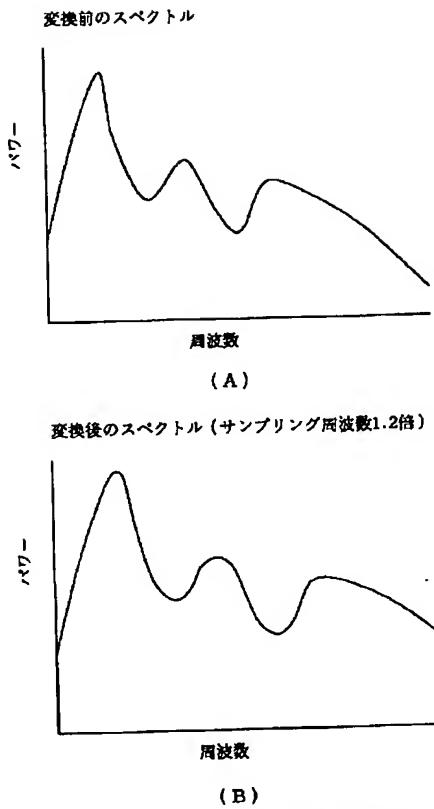


図2 D/A変換器の構成

【図3】

図3 D/A変換器のサンプリング周波数を変えることによる
音声スペクトルの変化の様子

【図4】

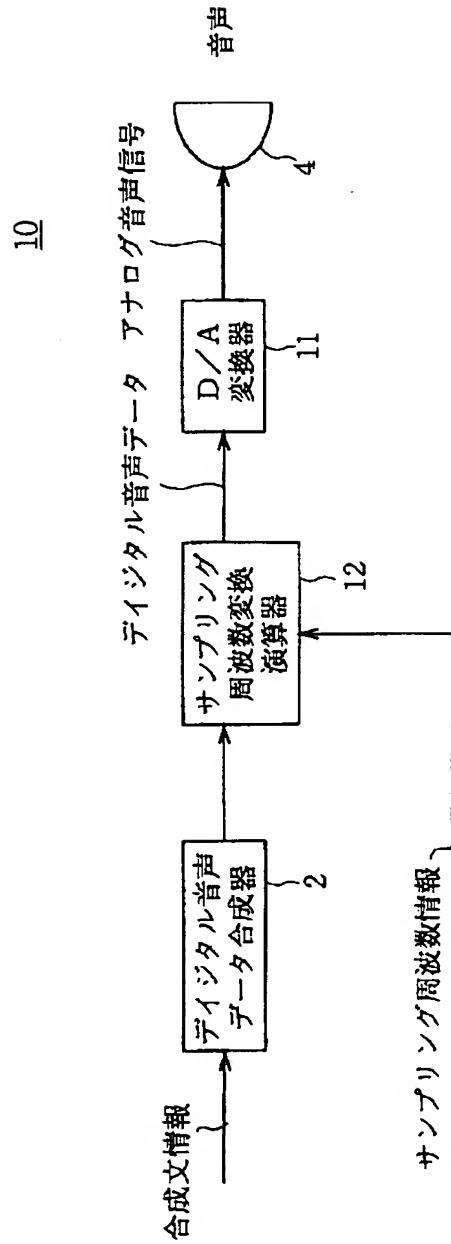


図4 第2実施例による音声合成装置の構成

THIS PAGE BLANK (USPTO)